AMAPS - a New Paradigm for Data Distribution and Analysis

Susan Paradise, Jet Propulsion Laboratory, California Institute of Technology
Amy Braverman, Jet Propulsion Laboratory, California Institute of Technology
Brian Wilson, Jet Propulsion Laboratory, California Institute of Technology

Technological advancements in computing power and storage have made possible the collection and analysis of massive scientific datasets. Distribution has improved dramatically in recent years by making the data web-accessible. However, the scientist interested in comparing or fusing data from multiple sources is faced with a true challenge: large datasets from different locations and in different data formats must be transferred to the analyst's machine, then co-located. The required order of events often makes global inter-comparison and analysis impractical.

Grid computing offers the opportunity to introduce a new paradigm, by combining data analysis and distribution in the same step. In a grid-enabled environment, datasets are virtual, with different parameters combined on the fly from different sources and reduced before transfer, according to the needs of the analyst. Such an environment is necessary to provide the breakthrough needed for global analysis of massive datasets, especially where multiple sources are required. In this talk, we demonstrate the Aerosol Measurement and Processing System (AMAPS) for data distribution and analysis of aerosol data using the above methodologies.